

Deterministic characterization of stochastic genetic circuits

Matthew Scott^{*†}, Terence Hwa^{*}, and Brian Ingalls[‡]

^{*}Center for Theoretical Biological Physics and Department of Physics, University of California at San Diego, La Jolla, CA 92093-0374; and [‡]Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1

Edited by Michael S. Waterman, University of Southern California, Los Angeles, CA, and approved March 12, 2007 (received for review November 26, 2006)

For cellular biochemical reaction systems where the numbers of molecules is small, significant noise is associated with chemical reaction events. This molecular noise can give rise to behavior that is very different from the predictions of deterministic rate equation models. Unfortunately, there are few analytic methods for examining the qualitative behavior of stochastic systems. Here we describe such a method that extends deterministic analysis to include leading-order corrections due to the molecular noise. The method allows the steady-state behavior of the stochastic model to be easily computed, facilitates the mapping of stability phase diagrams that include stochastic effects, and reveals how model parameters affect noise susceptibility in a manner not accessible to numerical simulation. By way of illustration we consider two genetic circuits: a bistable positive-feedback loop and a negative-feedback oscillator. We find in the positive feedback circuit that translational activation leads to a far more stable system than transcriptional control. Conversely, in a negative-feedback loop triggered by a positive-feedback switch, the stochasticity of transcriptional control is harnessed to generate reproducible oscillations.

intrinsic noise | phase diagram | synthetic biology | oscillations

Parallel advances in the conceptual understanding of gene regulation along with technological advances in molecular biology have given rise to the possibility of system-level quantitative kinetic measurements of living organisms (1) and synthetic genetic circuit designs (2, 3). Interpretation of time-series data from complex networks and reliable forward-design of gene circuits depend on detailed quantitative mathematical models (2, 4). These models generally take one of two largely exclusive forms: either deterministic formulations with reactant concentration varying continuously in time and governed by a system of rate equations or stochastic formulations that explicitly include the discrete and probabilistic change in reactant molecule numbers as each subsequent reaction occurs (5). Both approaches have benefits and associated limitations.

The great practical advantage of rate equation models is the ease with which the qualitative behavior of the system can be extracted. By focusing on the long-term behavior, the model dynamics are simplified, and one is able to gain insight into the expected response of the system (6). Rate equation models, however, neglect the fact that chemical reaction networks are composed of species that evolve on discrete space, jumping from some number of molecules to another as each reaction occurs (7). The resulting deviation from the deterministic formulation is called the “intrinsic noise” in the system (because the fluctuations arise from the reaction dynamics themselves and not from some external source) (8, 9). In cellular systems with small numbers of reactant molecules, the relative magnitude of the intrinsic noise can be large, and it can give rise to qualitatively different behavior than what rate equation models would predict. A system that has several possible stable states, for example, may be induced to spontaneous transitions between them as a result of intrinsic noise (10, 11), leading to a stochastic switching of states. In an excitable system, noise may cause oscillations to

occur in a model that is otherwise stable (12–14). With a given set of physical parameters, it is possible to simulate explicitly the individual chemical reaction events, including the effect of intrinsic noise (5). Nevertheless, the design of synthetic circuits, or therapeutics aimed at altering an existing network, require knowledge of the “phase diagram,” which involves a systematic mapping of the parameter space. There, stochastic simulation becomes prohibitively time-consuming even for reasonably simple genetic circuits involving two or three genes (see below), and analytical methods are needed.

A number of analytic studies have been done recently to model intrinsic noise in genetic circuits, much of them built on the linear noise approximation (15) and focused on the noise property itself, e.g., “noise propagation” through genetic networks (16, 17), the equilibrium distribution of fluctuations about multiple steady-states (18), and constructive effects of noise in signal processing (13, 19). There has been comparatively little work, however, aimed at providing tools to study the effect of intrinsic noise on the stability of systems where stochastic models exhibit qualitatively different behavior from their deterministic counterparts (20). Under these conditions, the linear noise approximation alone cannot predict qualitative changes in the observable dynamics of the system, as, for example, in the case of noise-induced oscillations (21). Here we present an analytic method, which we call the “effective stability approximation” (ESA), that extends the applicability of existing deterministic methods to include stochastic effects. The method is an extension of the linear noise approximation, including correction of stochasticity to the deterministic equations to the order $1/N$ (where N is the number of molecules in the system). It conveniently connects deterministic and stochastic descriptions, allowing for systematic exploration of parameter space while at the same time including the essential effect of intrinsic fluctuations. For the two model systems examined here, we find the ESA to capture reliably the essential features of those systems, correctly estimating the effect of intrinsic noise on the phase diagrams of systems dominated by as little as a few dozen molecules.

ESA can be applied to generic models of genetic circuits, and a brief tutorial is presented in *Methods* with the hope that the approach can be used by other investigators to include stochastic effects in deterministic models. The full mathematical details are presented in supporting information (SI). We illustrate the power of the method below by considering two examples: an autoregulator with positive feedback (an “autoactivator”) (22) and an excitable genetic oscillator linking positive and negative feedback loops (12, 23). The behavior of both circuits is conve-

Author contributions: M.S., T.H., and B.I. performed research and wrote the paper.

The authors declare no conflicts of interest.

This article is a PNAS Direct Submission.

Abbreviation: ESA, effective stability approximation.

[†]To whom correspondence should be addressed. E-mail: mscott@ctbp.ucsd.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0610468104/DC1.

© 2007 by The National Academy of Sciences of the USA

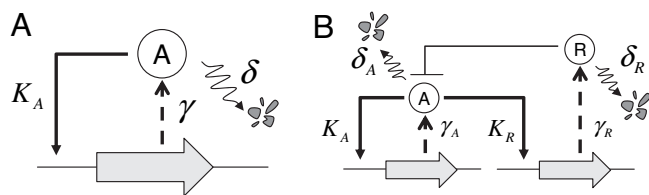


Fig. 1. Two example circuit motifs. (A) A positive-feedback loop capable of maintaining two stable states (22). (B) An excitable oscillator that exhibits noise-induced oscillations (12, 23). The autoactivator triggers the production of a repressor R that provides negative feedback control. Dashed arrows, lumped transcription and translation; bold filled arrows, activation; blunt arrow, repression; wavy arrows, degradation.

niently visualized by means of a phase diagram that cannot be practically constructed by using numerical simulations if stochastic effects are to be included. Furthermore, the analysis reveals that the system behavior is completely governed by a few dimensionless combinations of model parameters: combinations that would be very difficult to infer from simulation data alone. We hope that our presentation of the ESA method will make it accessible to modelers, bioengineers, and synthetic circuit designers for the analysis of various molecular circuits, while our description of the behaviors of the two model systems will provide quantitative-minded biologists with a concrete sense of the effect of stochasticity as well as a succinct means of characterization (e.g., a phase diagram with reduced variables).

Results and Discussion

Autoactivator. Perhaps the simplest circuit motif able to exhibit multiple stable states is the autoactivating positive-feedback loop (Fig. 14) (24). The circuit consists of a single gene encoding an activator. Several autoactivator circuits have been experimentally characterized, including the autoactivation of CI protein by the P_{RM} promoter of phage λ studied by Isaacs *et al.* (22) and the autoactivation of NtrC by the glnAp promoter of *Escherichia coli* studied by Atkinson *et al.* (23). The autoactivator circuit is expected to exhibit either a HIGH state, characterized by an elevated level of protein synthesis, or a LOW state, characterized by a low basal level of production. We simplify the model by assuming that the activator binding and mRNA turnover are fast compared with the lifetime of the protein activator. The effect of the activator is quantified by the “activation function” $g(A/K_A, f)$, where A is the activator concentration, K_A is the equilibrium dissociation constant of the activator and its cognate binding site, and f is the maximum fold-activation in the circuit. As a particular example, we assume a Hill-form for the activation function $g(A/K_A, f)$.

$$g\left(\frac{A}{K_A}, f\right) = \frac{f^{-1} + \left(\frac{A}{K_A}\right)^n}{1 + \left(\frac{A}{K_A}\right)^n}, \quad [1]$$

with cooperative activation ($n = 2$) (4). The resulting model is a single kinetic equation governing the activator concentration $A(t)$ (22, 25) (Fig. 1A),

$$\frac{dA}{dt} = \gamma g(A) - \delta \cdot A, \quad [2]$$

where γ is the fully activated rate of protein synthesis and δ is the protein degradation rate (which in prokaryotes is often estimated from the growth rate due to growth-mediated dilution).

In the deterministic limit, when the number of reactant molecules is very large, we expect Eq. 2 to adequately describe the system behavior. Once initial transients have died out, the system will approach a steady state, and A reaches its steady-state value A_s where the rate of synthesis and degradation balance, i.e., $\gamma g(A_s) = \delta A_s$. The stability of the steady state is determined by the response of the system to a small perturbation A_p , found by linearizing Eq. 2 about A_s ,

$$\frac{dA_p}{dt} = [\gamma g'(A_s) - \delta] \cdot A_p \equiv \lambda \cdot A_p. \quad [3]$$

The expression in the square brackets $\lambda \equiv [\gamma g'(A_s) - \delta]$ is a constant that depends on the model parameters. If λ is positive, the small perturbations will grow in time (A_s is an “unstable state”), whereas if λ is negative, the small perturbation will decay (A_s is a “stable state”). In the stable case, the long-term state of the system can be thought of as a point located at the bottom of a valley (or basin of attraction): the more negative the constant λ , the steeper the valley. As the model parameters are varied, the valley may become more flat ($\lambda \approx 0$) or even develop into a mountain ($\lambda > 0$), resulting in a loss of stability. The parameter space is divided into regions of different qualitative behavior (as in Fig. 24, black curve); the threshold between these domains indicates where λ has changed sign and is called the “phase boundary.” Although the model seems to depend on a large family of parameters (γ , δ , K_A , etc.), the stability of the deterministic model is actually described by two dimensionless combinations of these parameters: the ratio of the protein concentration with fully activated promoter ($A_0 = \gamma/\delta$) to the dissociation constant, A_0/K_A , and the fold-activation, f .

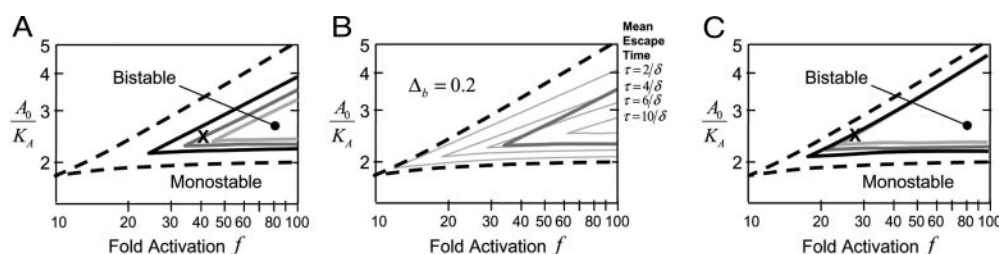


Fig. 2. Stability phase plot for the autoactivator (Fig. 1A), including the effect of intrinsic noise. (A) The black dashed curve is the phase boundary of the deterministic model with transcriptional activation (A_0/K_A is the fully activated protein concentration scaled by the activator/DNA dissociation constant). Increasing the level of intrinsic noise by increasing the discreteness parameter Δ_b (i.e., increasing the “burstiness” of translation or decreasing the number of molecules) diminishes the parameter regime of reliable bistability ($Re[\lambda'] < 0$). Here, $\Delta_b = 0.1$ (black solid line), 0.2 (dark gray), and 0.3 (light gray). (B) The average escape time from the stable state is an indicator of the permanence of the bistability. Here, the dark gray curve from A corresponds to an escape time of approximately $\tau = 6$, where time has been scaled relative to the protein lifetime δ^{-1} . (C) As in A, but now with translational activation. The range of bistability is considerably widened as transitions from the LOW to the HIGH state are suppressed. Here, $K_A V_{\text{cell}} = 25$ molecules, and the fully activated burst size is $b = 4$ (black), 9 (dark gray), and 14 (light gray).

The ESA we propose is an approximation that allows the average effect of intrinsic noise to be expressed as a positive correction to λ ,

$$\lambda' = \lambda + \lambda_{\text{corr}} (\lambda_{\text{corr}} > 0), \quad [4]$$

(see Eq. 13 below). The correction reflects an effective “flattening” of the local landscape by stochastic fluctuations, making it easier for the system to escape from the basin of attraction. Adopting this perspective allows the analysis used to study the deterministic model to be extended to the stochastic model with only minor modification. With λ' corrected to include the effect of the intrinsic noise, the new phase boundaries are drawn to coincide with points in parameter space where $\lambda' = 0$.

A major source of intrinsic noise in gene regulatory networks is so-called translational bursting (7, 26), where each mRNA transcript is translated into several peptides before the message is degraded, leading to a “burst” of protein synthesis. Typical values of the “burst size” b can vary from close to zero for poorly translated genes (27) up to several dozen (28, 29) depending on the rate of translation and the lifetime of the transcript. When intrinsic noise is included in the autoactivator model, and the procedure described in detail in Section IIIA of SI is applied, we find the correction to λ is $\lambda_{\text{corr}} \propto \Delta_b/\lambda^2$, where

$$\Delta_b = \frac{(b+1)}{2} \frac{1}{K_A V_{\text{cell}}} = \frac{(b+1)}{2} \frac{1}{N_A}, \quad [5]$$

is a third dimensionless quantity we call the “discreteness parameter.” This parameter captures the average change in protein number when a synthesis or degradation event occurs, scaled relative to the protein number required to initiate activation $N_A = K_A \times V_{\text{cell}}$, where K_A is the activator dissociation constant and V_{cell} is the cell volume. Increasing the discreteness parameter Δ_b increases the magnitude of the discrete change in activator numbers and, therefore, increases the relative magnitude of the perturbation to the system caused by the intrinsic noise. One would expect the circuit to switch more readily from stable state to stable state as the magnitude of the intrinsic noise is increased, thereby reducing the average stability of the circuit. On the other hand, as the number of activator molecules increases ($N_A \rightarrow \infty$), the discreteness parameter vanishes and the behavior of the system is fully described by the deterministic model. Thus, the discreteness parameter Δ_b represents a distillation of the complicated effect of intrinsic noise on the model behavior, captured in a compact expression that would be difficult to extract from numerical simulation data.

As shown in Fig. 2A, for the autoactivator the parameter space is divided into regions of bistability (two stable states) and monostability (one stable state). The bistability is most easily lost near the phase boundary separating the bistable and monostable states. The circuit parameters of Isaacs *et al.* (22) lie close to the left-hand tip of the black triangle in Fig. 2A ($f \approx 10$), and as they observed in their experiments, the noise overwhelms bistability in such a system (cf. figure 2A of ref. 22), leading to rapid transitions between the stable states. A much greater fold-activation is required to maintain two distinct stable states (as likewise noted by the authors).

Actually, once noise is allowed in the autoactivator model, one no longer has stability in the strictest sense because there is always a chance that a perturbation will switch the system from one steady state to the other. With noise, it is not a question of stability, but rather the average escape time from the steady state (10, 11). The longer the escape time (compared with other time scales in the problem), the more “stable” the system. To emphasize the effect of the intrinsic noise on the stability phase plot, we consider a system with a small number of activator proteins ($K_A V_{\text{cell}} = 25$ molecules). Using the parameters $\gamma = 2$

protein \cdot min $^{-1}$, $\delta^{-1} = 30$ min (a half-life of ≈ 20 min), $K_A = 25$ nM, and a burst size of $b = 10$, the discreteness parameter in *E. coli* ($V_{\text{cell}} \approx 1 \mu\text{m}^3$) is $\Delta_b \approx 0.2$. From Fig. 2A (dark gray curve), a maximum fold-activation of $f \geq 40$ is necessary to ensure long-lived bistable states (shown as a cross on the plot). It is possible to explicitly compute the average escape time from the stable states for this simple model (see refs. 30 and 31 and Section II of SI). Fig. 2B compares the average escape time as a function of A_0/K_A and f for the case above, with $\Delta_b = 0.2$ (dark gray curve in Fig. 2A). Along the dark gray curve, the escape time is $\tau = (3 \pm 0.5)$ h, which is approximately six times longer than the protein lifetime (which sets the basic time scale of the system’s “memory”).

The escape time is an indirect measure of the system’s stability. We have developed a more direct method that measures the effective rate of divergence of an ensemble of stochastic trajectories. This method is of general applicability and allows a direct evaluation of the accuracy of the ESA. The details of that calculation are reserved for SI (see Section III-A.2). Comparing λ' with the effective rate of divergence in the stochastic simulations of the autoactivator, the ESA is found to be accurate for systems with $\Delta_b \leq 0.25$.

The burst size b can be reduced by decreasing the rate of translation, and indeed Ozbudak *et al.* (27) suggest that many poorly translated genes in *E. coli* could be the result of evolutionary selection against burst noise. Alternatively, the method of control in the circuit can be shifted from transcriptional to translational activation. Although the simple deterministic model remains unchanged for either choice of transcriptional or translational control, the resulting stochastic model exhibits improved stability for translational activation.

Fig. 2C shows the result of putting the translation rate under control of the activator. Decreasing the translation rate in the LOW state has the effect of shifting the upper branch of the phase boundary, indicating a decrease in transitions from the LOW to the HIGH state. The translational autoactivator can tolerate a larger range of transcription rates (i.e., higher γ) and a lower maximum fold-activation ($f \geq 20$), even for large burst size. As above, with $\gamma = 2$ protein \cdot min $^{-1}$, $\delta^{-1} = 30$ min, $K_A = 25$ nM, $V_{\text{cell}} = 1 \mu\text{m}^3$ and a fully activated burst size $b = 10$, a fold-activation of $f \geq 25$ is required to sustain the bistability (shown as a cross on the plot), almost half that required in the transcriptional autoactivator above.

To generate the phase plot for a given stochastic model requires division of the parameter space of the model (Eq. 2) into a fine grid, with stochastic simulation performed at each point. Even after several such simulations are generated, it is unlikely that the discreteness parameter Δ_b will suggest itself as a key measure of the magnitude of intrinsic noise. The ESA method provides not only a rapid overview of the parameter space, but also compact expressions characterizing the effect of intrinsic noise on the observable dynamics. In the next section, we apply ESA to the analysis of a more elaborate circuit model.

Genetic Oscillator. Oscillating systems underlie many physiological processes in the cell, from circadian rhythms (32) to the cell cycle itself (33). In addition to the natural systems, several synthetic genetic oscillator designs have been studied, including the mutually repressing ring-oscillator (*Repressilator*) of Elowitz and Leibler (34) and the activator–repressor design of Atkinson and coworkers (23) (which has a great deal in common with the model discussed below). A recurring motif in experimentally characterized networks is a negative-feedback loop serving as a system reset (32, 35). Without some time delay or intervening mechanism to prevent reversibility, the system will rapidly approach an intermediate equilibrium, and it is found both theoretically (36) and experimentally (33) that a negative feedback loop alone is not sufficient to maintain reliable oscillations.

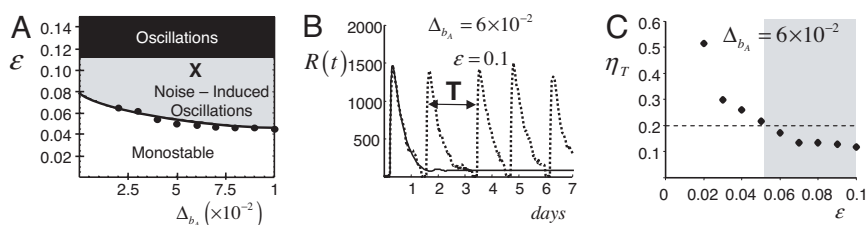


Fig. 3. Noise induced oscillations in the excitable oscillator. (A) Stability phase plot as a function of the scaled repressor degradation rate $\varepsilon = \delta_R/\delta_A$ for the circuit shown in Fig. 1B. The discreteness in the activator synthesis, Δ_{b_A} , characterizes the average discrete change in activator concentration during each reaction, and consequently the magnitude of the intrinsic noise. The intrinsic noise expands the region of instability (gray), extending the parameter range over which oscillations are expected to occur. The deterministic phase boundary is located at $\varepsilon \approx 0.12$ (dashed line separating the black and gray regions). The solid line is the phase boundary predicted from the roots of Eq. 12, and filled circles denote the phase boundary found by stochastic simulation (see text). The model and parameters are as in Vilar *et al.* (12). (B) The circuit exhibits noise-induced oscillations (dotted line) with interspike time T . The parameters used in the simulation correspond to a deterministically stable system (solid line). Numerical simulation data were generated by using Gillespie's direct method (5), with parameters as used in ref. 12 and $\varepsilon = 0.1$, $\Delta_{b_A} = 6 \times 10^{-2}$ (cross in A). (See Section III-C of SI.) (C) A plot of the noise-to-signal ratio $\eta_T = (\langle T \rangle - T)^2/12/T$ as a function of ε . The oscillations are regular when η_T is small (the region of noise-induced oscillations predicted by the ESA is gray), and η_T was calculated by using at least 200 spikes for each point.

If, however, the feedback repressor is controlled by a bistable autoactivator, the oscillations become more robust and coherent because the bistable switch acts as a ratchet that “locks” into the HIGH state generating a large amount of repressor to feedback and reset the system to the LOW state where the system remains until the activator accumulates over a critical threshold to initiate another cycle (37). This motif is highly represented in natural gene networks (35), and we will use the ESA to ascertain the contribution of intrinsic noise to the performance of such an oscillator.

We consider the generic model proposed by Vilar *et al.* (12) to describe circadian rhythms in eukaryotes, with a transcriptional autoactivator driving expression of a repressor that provides negative control by sequestering activator proteins through dimerization (13, 38). The repressor and activator form an inert complex until the activator degrades, recycling repressor back into the system. In their model, the degradation rate of the activator, δ_A , is the same irrespective of whether it is bound in the inert complex or free in solution. We simplify their original model somewhat, and as in the previous section, we assume fast activator/DNA binding and rapid mRNA turnover, leading to a reduced set of rate equations governing the concentration of activator A , repressor R , and the inert dimer C ,

$$\begin{aligned}\frac{dA}{dt} &= \gamma_A \mathcal{G}\left(\frac{A}{K_A}, f_A\right) - \delta_A \cdot A - \kappa_C \cdot A \cdot R \\ \frac{dR}{dt} &= \gamma_R \mathcal{G}\left(\frac{A}{K_R}, f_R\right) - \delta_R \cdot R - \kappa_C \cdot A \cdot R + \delta_A \cdot C \\ \frac{dC}{dt} &= \kappa_C \cdot A \cdot R - \delta_A \cdot C.\end{aligned}\quad [6]$$

We further assume no cooperativity in activator binding ($n = 1$ in the activation function g) and the nominal parameter set used in ref. 12. For this more complicated system, there is a larger number of dimensionless combinations of parameters that characterize the system dynamics. The scaled repressor degradation rate $\epsilon = \delta_R/\delta_A$ is a key control parameter in the model because oscillations occur in the deterministic system only for an intermediate range of this parameter. For the nominal parameter set used in ref. 12, the deterministic model exhibits oscillations over the range $0.12 < \epsilon < 40$ (Fig. 3A, black region). We will focus on the parameter regime near to the phase boundary at $\epsilon \approx 0.12$ and examine the role intrinsic noise plays in generating regular oscillations from a deterministically stable system.

Applying the ESA to the oscillator model, the parameter $\Delta_{b_A} = (b_A + 1)/(2 \cdot K_A \cdot V_{\text{cell}})$ emerges as an important measure quantifying the discreteness in activator synthesis (see Eq. 36 in SI).

Here again, b_A is the burst size in the activator synthesis, K_A is the activator/DNA dissociation constant, and V_{cell} is the cell volume. (Here, $V_{\text{cell}} = 100 \mu\text{m}^3$ as is appropriate for eukaryotic cells.)

Using the nominal parameter set of Vilar *et al.* (12) in our reduced model leads to a burstiness in activator synthesis of $b_A = 5$ (giving $\Delta_{b_A} = 6 \times 10^{-2}$) and a burstiness in repressor synthesis of $b_R = 10$. The phase boundary predicted by the ESA is shown as a solid line in Fig. 3A, bounding a region of parameter space between the deterministic phase boundary where qualitatively different behavior is expected from the stochastic model. We examine the system behavior in this region by running a stochastic simulation using the parameter choice $\varepsilon = 0.1$ and $\Delta_{b_A} = 6 \times 10^{-2}$ (denoted by a cross in Fig. 3A). With this choice, the deterministic model is stable (Fig. 3B, black line). Nevertheless, a stochastic simulation of the same model, including protein bursting and stochastic dimerization, clearly shows oscillations (Fig. 3B, dotted line).

The time between successive peaks in the stochastic simulation of Fig. 3B is denoted by T . As is clear from Fig. 3B, T is itself a random variable. Each simulation run generates a collection of interspike times from which the mean $\langle T \rangle$ and the variance $\langle (T - \langle T \rangle)^2 \rangle$ can be calculated. Following Steuer *et al.* (13), the quality of the noise-induced oscillations is measured by using the noise-to-signal ratio $\eta_T = \langle (T - \langle T \rangle)^2 \rangle^{1/2} / \langle T \rangle$, and the system is said to exhibit regular oscillations where η_T is small (13, 38). The dependence of η_T on the repressor degradation rate ε is shown in Fig. 3C, with the discreteness parameter $\Delta_{b_A} = 6 \times 10^{-2}$ (as in Fig. 3B), using at least 200 spikes to calculate η_T . At low repressor degradation rate, the noise-to-signal ratio is high, indicating large variance in the interspike time T and corresponding to a stable (i.e., nonoscillatory) system. As the repressor degradation rate is increased, the variance in the interspike time T decreases with a consequent decrease in the noise-to-signal ratio η_T , indicative of a more regularly oscillating system. Physically, the intrinsic noise in this parameter range is sufficient to drive the system away from the deterministically stable steady state, yet the noise is not so strong that the return trajectory through phase space is much affected.

As in the autoactivator model, it is useful to compare the phase boundary predicted by the ESA to some independent measure of stability, in this case η_T . In Fig. 3C, the ESA phase boundary (for $\Delta_{b_A} = 6 \times 10^{-2}$) is denoted by the interface between the white and gray regions, corresponding to a value of $\eta_T \approx 0.2$. Using data such as that shown in Fig. 3C, the points in the phase plot with $\eta_T = 0.2$ can be found for a range of discreteness parameter Δ_{b_A} (Fig. 3A, filled circles). These points correspond very well to the phase boundary calculated by using the ESA

(Fig. 3A, solid line). The results are as one would expect: near the deterministic phase boundary, very little molecular noise is required to sustain oscillations, and reasonable periodicity persists even for small values of the discreteness parameter ($\Delta_{b_A} \rightarrow 0$, $b_R \neq 0$). As the repressor degradation rate ε is decreased to a region favoring stability, more noise is required to overcome the deterministic stability of the system and initiate the autoactivator trigger. It is illustrative to remark that each data point in Fig. 3A, obtained from stochastic simulation (5), took ≈ 1 day to generate on a dual-processor desktop computer because at low repressor degradation rate, a large separation of timescales is introduced, necessitating long stochastic simulation runs to capture the slowly varying dynamics of the system. By contrast, the solid line generated from the roots of Eq. 12 took < 1 h to produce on the same machine. Thus, even for a two-gene circuit with several degrees of freedom, the ESA affords a compact and convenient means to survey the phase space, drawing attention to those regions of particular interest that may be probed in more detail by more realistic (although also more computationally costly) stochastic simulation methods.

Methods

The ESA can be applied to generic models of genetic circuits in a straightforward way. Here, a brief outline of the method is provided. A self-contained tutorial on stochastic modeling and the ESA is provided in SI.

A useful abstraction of genetic regulatory networks is as a system of ordinary differential equations (39, 40). (Here, and throughout, we shall assume a spatially homogeneous environment.) We denote the concentrations of the reactants of interest by the state vector \mathbf{x} , where the x_i values correspond to the concentration of mRNA, transcription factors, protein products, etc. The kinetic equation governing the evolution of the system takes the form $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x})$, where \mathbf{f} is a vector of nonlinear functions of the state variables. We can estimate the long-time, or steady-state, behavior of the model by first computing the equilibrium points \mathbf{x}_s that satisfy the algebraic constraint $\mathbf{f}(\mathbf{x}_s) = \mathbf{0}$. We then Taylor expand the reaction rate vector about the equilibrium point by making the substitution $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_p$ (where \mathbf{x}_p is an infinitesimal perturbation away from \mathbf{x}_s), and retain only linear terms in \mathbf{x}_p . The resulting dynamics of \mathbf{x}_p are given by $(d/dt)\mathbf{x}_p = \mathbf{J}\mathbf{x}_p$, where \mathbf{J} is the Jacobian or response matrix: $J_{ij} = \partial f_i / \partial x_j$. The eigenvalues of \mathbf{J} are the matrix analogue of the parameter λ introduced in Eq. 3, and in a similar fashion if the eigenvalues all have negative real-part, then \mathbf{x}_s is a stable steady state. [There are, of course, limitations to how far one can trust the linearization (6), but for our purposes it is sufficient as a first approximation.]

To include stochastic effects in the mathematical model, chemical reaction rates must be rewritten in terms of the reaction “propensity” and “stoichiometry” (5). For example, in the positive autoactivator example above, with the individual synthesis and degradation stoichiometries written explicitly, the deterministic model equations (Eq. 2) read

$$\begin{aligned} \text{bursty synthesis: } & A \xrightarrow{\nu_1} A + b; \quad \nu_1 = \frac{\gamma}{b}g(A), \\ \text{linear degradation: } & A \xrightarrow{\nu_2} A - 1; \quad \nu_2 = \delta A. \end{aligned} \quad [7]$$

We encode this information concisely as the “propensity vector” $\nu = [\nu_1, \nu_2] = [\gamma g(A)/b, \delta A]$ and the stoichiometry matrix $\mathbf{S} = [b, -1]$. The discrete change in molecule numbers after the completion of a chemical reaction causes a deviation from the deterministic solution because the deterministic model assumes

an infinitesimally small and continuous change in the state. (Consequently, the deterministic model only applies to systems with large numbers of molecules.) We denote the deviation of the stochastic model from the deterministic model by the fluctuating quantity $\omega\alpha(t)$, where $\omega = 1/\sqrt{V_{\text{cell}}}$ and $\alpha(t)$ describes the stochastic deviation in each species \mathbf{x} . The $\sqrt{V_{\text{cell}}}$ scaling arises from the observation that the relative magnitude of the intrinsic noise scales roughly as the inverse square root of the number of molecules (15). Elf and Ehrenberg (21) developed an algorithmic expression for the statistics of α using the linear noise approximation of vanKampen (15). In that formulation, the mean and covariance of the fluctuations about the deterministic state are written compactly in terms of the propensity vector ν and the stoichiometry matrix \mathbf{S} ; here, we shall apply their method to characterize the fluctuations about the stable state. The first step in the calculation of the moments of the fluctuations $\omega\alpha(t)$ is to construct the auxiliary matrices $\mathbf{\Gamma}$ and \mathbf{D} , evaluated at the stable state \mathbf{x}_s ,

$$\Gamma_{ij}(t) = \frac{\partial[\mathbf{S}\cdot\nu]_i}{\partial x_j} = \frac{\partial f_i}{\partial x_j} \quad \mathbf{D} = \mathbf{S}\cdot\text{diag}[\nu]\cdot\mathbf{S}^T. \quad [8]$$

The drift matrix $\mathbf{\Gamma} = \mathbf{J}$ is the response matrix (or Jacobian) described above and reflects the local stability of the deterministic system to small perturbations (41). The diffusion matrix \mathbf{D} captures the strength of the fluctuations and is related to the magnitude of the reaction step size (21, 42). It is straightforward to show that to leading order in ω the mean of the fluctuations is zero ($\langle\alpha\rangle = 0$) and the variance, denoted by the symmetric matrix $\Xi = \langle\alpha\alpha^T\rangle$, is determined by the solution of the system of algebraic equations (15), $\mathbf{\Gamma}\cdot\Xi + \Xi\cdot\mathbf{\Gamma}^T + \mathbf{D} = \mathbf{0}$. Because the fluctuations about the stable state are stationary, the time autocorrelation function depends on the time difference only and is given by the matrix exponential

$$\langle\alpha(t)\alpha^T(t-\tau)\rangle = \exp[\mathbf{\Gamma}\tau]\cdot\Xi. \quad [9]$$

The effect of the fluctuations on the deterministic steady-state is calculated by including an additional term in the deterministic linearization above: $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_p + \omega\alpha$. Linearizing \mathbf{J} in ω , we have a stochastic differential equation governing the decay of the perturbation modes \mathbf{x}_p

$$\frac{d}{dt}\mathbf{x}_p = [\mathbf{J}^{(0)} + \omega\mathbf{J}^{(1)}(t)]\mathbf{x}_p. \quad [10]$$

The fluctuations affect the decay of the infinitesimal disturbance \mathbf{x}_p as well as the dynamics of the average $\langle\mathbf{x}_p\rangle$, which (provided $\omega\mathbf{J}^{(1)}(t) \ll \mathbf{J}^{(0)}$) is approximately governed by the convolution equation (43, 44)

$$\frac{d}{dt}\langle\mathbf{x}_p(t)\rangle = \mathbf{J}^{(0)}\langle\mathbf{x}_p(t)\rangle + \omega^2 \int_0^t \mathbf{J}_c(t-\tau)\langle\mathbf{x}_p(\tau)\rangle d\tau, \quad [11]$$

where $\mathbf{J}_c(t-\tau) = \langle\mathbf{J}^{(1)}(t)e^{\mathbf{J}^{(0)}(t-\tau)}\mathbf{J}^{(1)}(\tau)\rangle$ is made up of linear combinations of the cross-correlations $\langle\alpha_i(t)\alpha_j(\tau)\rangle$ given by the i th row and the j th column of the right-hand side of Eq. 9. In the noiseless case, the stability of the perturbation \mathbf{x}_p is determined by the eigenvalues of $\mathbf{J}^{(0)}$: $\text{diag}\{\lambda_i\} = \mathbf{P}^{-1}\mathbf{J}^{(0)}\mathbf{P}$ where the matrix \mathbf{P} is made of the eigenvectors of $\mathbf{J}^{(0)}$. The analogues of the eigenvalues for the convolution equation above are found from the poles of the Laplace transform, denoted λ' , which solve the resolvent equation (45)

$$\det \left[\lambda' \mathbf{I} - \mathbf{J}^{(0)} - \frac{1}{V_{\text{cell}}} \hat{\mathbf{J}}_c(\lambda') \right] = 0. \quad [12]$$

Here ω^2 has been replaced by V_{cell}^{-1} and $\hat{\mathbf{J}}_c(s) = \int_0^\infty \mathbf{J}_c(t) e^{-st} dt$ is the Laplace transform of $\mathbf{J}_c(t)$. If the deterministic eigenvalues are distinct, we can further approximate the effective eigenvalue λ'_i by

$$\lambda'_i = \lambda_i + \frac{1}{V_{\text{cell}}} [\mathbf{P}^{-1} \cdot \hat{\mathbf{J}}_c(\lambda_i) \cdot \mathbf{P}]_{ii}, \quad [13]$$

where $[\cdot]_{ii}$ denotes the i th diagonal entry of the matrix. Physically, we interpret the leading-order noise correction as the power in the fluctuations at eigenfrequency λ_i projected in the eigendirection of λ_i . Because the correction term is quadratic, it is always positive and thus destabilizes the eigenmode on which it is projected. (Hence, in Eq. 4 we write $\lambda_{\text{corr}} > 0$.)

It often happens that out of the term $1/V_{\text{cell}} [\mathbf{P}^{-1} \cdot \hat{\mathbf{J}}_c(\lambda_i) \cdot \mathbf{P}]_{ii}$ there appears a small parameter that quantifies the effect of the

intrinsic noise. (For the two examples above, the small parameters are Δ_b and Δ_{b_A} , each characterizing the discreteness of the protein change.) In the limit that this parameter goes to zero, the effect of the intrinsic noise becomes negligible, at least in that particular eigenmode.

Finally, the ESA can be easily implemented in a symbolic computational environment, without attending to the mathematical details (see Section IV of SI). A version of the ESA coded in Mathematica is freely available from the authors upon request.

We thank Jian Liu, Francis Poulin, and Stefan Klumpp for critical reading and constructive comments on the manuscript. M.S. was supported by a postdoctoral fellowship from Canada's Natural Sciences and Engineering Research Council (NSERC). B.I. is supported by an NSERC Discovery grant. This work was supported in part by National Science Foundation Grant MCB0417721 (to T.H.) and by the Physics Frontiers Center-sponsored Center for Theoretical Biological Physics Grants PHY-0216576 and PHY-0225630.

1. Kobilier O, Rokney A, Friedman N, Court DL, Stavans J, Oppenheim AB (2005) *Proc Natl Acad Sci USA* 102:4470–4475.
2. Hasty J, McMillen D, Collins JJ (2002) *Nature* 420:224–230.
3. Kaern M, Blake WJ, Collins JJ (2003) *Ann Rev Biomed Eng* 5:179–206.
4. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R (2005) *Curr Opin Gen Dev* 15:116–124.
5. Gillespie DT (1977) *J Phys Chem* 81:2340–2361.
6. Strogatz SH (1994) *Nonlinear Dynamics and Chaos* (Westview-Perseus, New York).
7. Kaern M, Elston TC, Blake WJ, Collins JJ (2005) *Nat Rev Genet* 6:451–464.
8. Swain PS, Elowitz MB, Siggia ED (2002) *Proc Natl Acad Sci USA* 99:12795–12800.
9. vanKampen NG (1992) *Stochastic Process in Physics and Chemistry* (North-Holland-Elsevier, Amsterdam).
10. Aurell E, Sneppen K (2002) *Phys Rev Lett* 88:048101.
11. Walczak AM, Onuchic JN, Wolynes PG (2005) *Proc Natl Acad Sci USA* 102:18926–18931.
12. Vilar JMG, Kueh HY, Barkai N, Leibler S (2002) *Proc Natl Acad Sci USA* 99:5988–5992.
13. Steuer R, Zhou C, Kurths J (2003) *BioSystems* 72:241–251.
14. Suel GM, Garcia-Ojalvo J, Liberman LM, Elowitz M (2006) *Nature* 440:545–550.
15. vanKampen NG (1976) *Adv Chem Phys* 34:245–308.
16. Tanase-Nicola S, Warren PB, tenWolde PR (2006) *Phys Rev Lett* 97:068102.
17. Pedraza JM, vanOudenaarden A (2005) *Science* 307:1965–1969.
18. Tomioka R, Kimura H, Kobayashi TJ, Aihara K (2004) *J Theor Biol* 229:501–521.
19. Paulsson J, Berg OG, Ehrenberg M (2000) *Proc Natl Acad Sci USA* 97:7148–7153.
20. DeVille REL, Muratov CB, Vanden-Eijnden E (2006) *J Chem Phys* 124:231102.
21. Elf J, Ehrenberg M (2003) *Genome Res* 13:2475–2484.
22. Isaacs FJ, Hasty J, Cantor CR, Collins JJ (2003) *Proc Natl Acad Sci USA* 100:7714–7719.
23. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ (2003) *Cell* 113:597–607.
24. Ferrell JE, Jr (2002) *Curr Opin Cell Biol* 14:140–148.
25. Keller AD (1995) *J Theor Biol* 172:169–185.
26. Thattai M, vanOudenaarden A (2001) *Proc Natl Acad Sci USA* 98:8614–8619.
27. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, vanOudenaarden A (2002) *Nat Genet* 31:69–73.
28. Kennell D, Riezman H (1977) *J Mol Biol* 114:1–21.
29. Cai L, Friedman N, Xie XS (2006) *Nature* 440:358–362.
30. Gardiner CW (2004) *Handbook of Stochastic Methods* (Springer, New York), 3rd Ed.
31. Kepler TB, Elston TC (2001) *Biophys J* 81:3116–3136.
32. Goldbeter A (1997) *Biochemical Oscillations and Cellular Rhythms* (Cambridge Univ Press, Cambridge, UK).
33. Pomerening JR, Kim SY, Ferrell JE, Jr (2005) *Cell* 122:565–578.
34. Elowitz MB, Leibler S (2000) *Nature* 403(6767):335–338.
35. Dunlap JC (1999) *Cell* 96:271–290.
36. Pomerening JR, Sontag ED, Ferrell JE, Jr (2003) *Nat Cell Biol* 5:346–351.
37. Cross FR, Siggia E (2005) *Dev Cell* 9:309–310.
38. Guantes R, Poyatos JF (2006) *PLoS Comp Biol* 2:188.
39. Conrad ED, Tyson JJ (2006) in *System Modeling in Cellular Biology*, eds Szallasi Z, Stelling J, Periwál V (MIT Press, Cambridge, MA), Chap 6, pp 97–124.
40. Kaern M, Weiss R (2006) in *System Modeling in Cellular Biology*, eds Szallasi Z, Stelling J, Periwál V (MIT Press, Cambridge, MA), Chap 13, pp 269–296.
41. Ali F, Menzinger M (1999) *Chaos* 9:348–356.
42. Scott M, Ingalls B, Kaern M (2006) *Chaos* 16:026107.
43. Bourret RC (1965) *Can J Phys* 43:619–639.
44. vanKampen NG (1976) *Phys Rep* 24:171–228.
45. Grossman SI, Miller RK (1973) *J Diff Eq* 13:551–566.